

# Can Automated Responses to Cyber Incidents be supported by Artificial Intelligence?

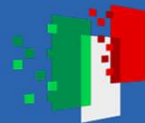
Manuel Zambelli, PhD student in Advanced-Systems Engineering  
Prof. Barbara Russo, Supervisor and Coordinator of the PhD program  
Software Engineering and Autonomous Systems (SEAS)  
Faculty of Engineering, Free University of Bolzano



Finanziato  
dall'Unione europea  
NextGenerationEU



Ministero  
dell'Università  
e della Ricerca



Italiadomani  
PIANO NAZIONALE  
DI RIPRESA E RESILIENZA



# Outline

- Motivation(s)
- Log files
- Security Information and Event Management (SIEM)
- Security Orchestration, Automation, and Response (SOAR)
- Manual labelling for AI Integration
- Incident descriptions
- Generation of the incident description as a vector
- Conclusions

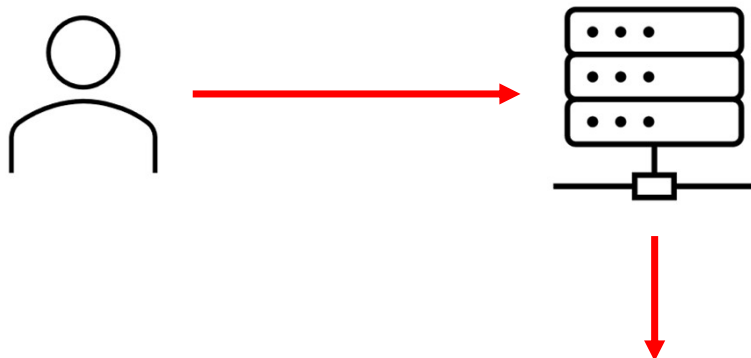
# Motivation 1: False Positives

- Problem of **False Positives** (FP) in Machine Learning
- Cyber incidents which are **misclassified**
- **Increases the costs** in the company (more working hours)
- Examples:
  - User login after failed attempts
  - Unauthorized access which does not compromise the system or steal data

## Motivation 2: Cyber Taxonomy

- **NIS2** (EU Directive 2022/2555, D.Lgs. n.138/2024)
- Starting from January 2026: communication of cyber incidents with a significant impact to the **Computer Security Incident Response Team (CSIRT)**
- The **Agenzia per la Cybersicurezza Nazionale (ACN)** defines a taxonomy (**TC-ACN**) for incident notifications:  
4 macrocategories, 22 predicates, 144 values

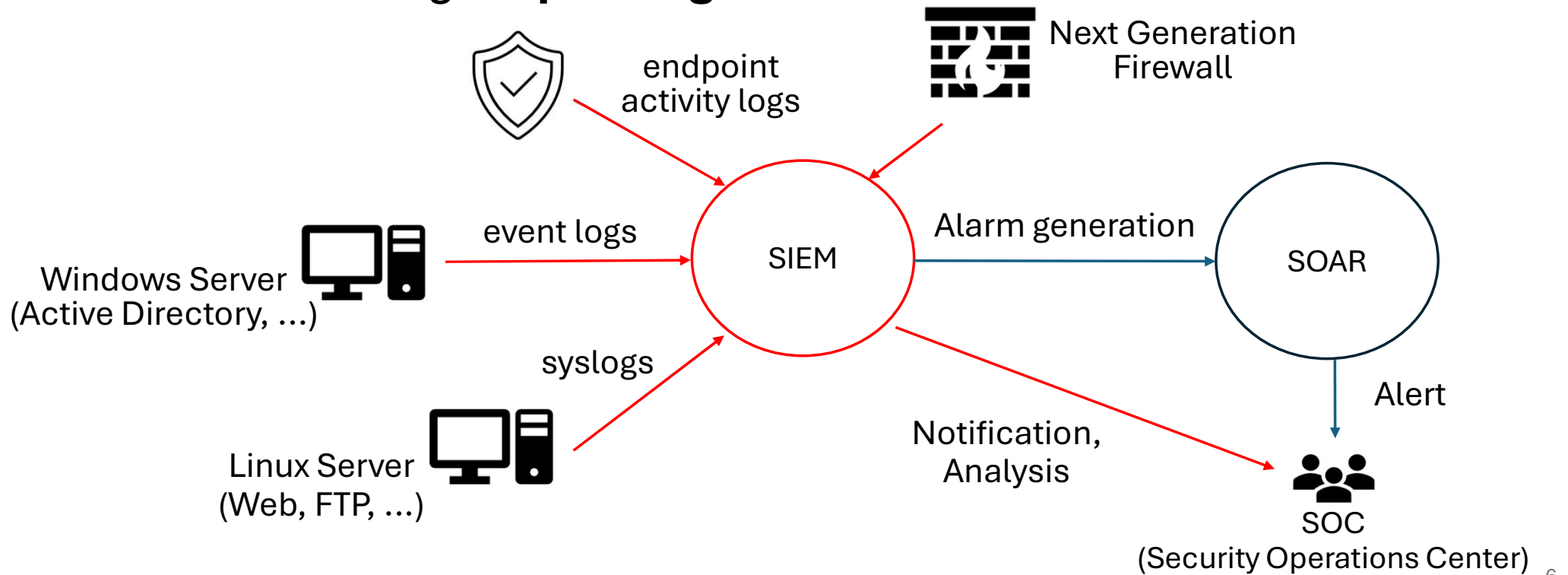
# Log Generation in Network Devices



|                                                                                     |                         |                     |      |                           |
|-------------------------------------------------------------------------------------|-------------------------|---------------------|------|---------------------------|
|    | Überwachung erfolgreich | 02.11.2025 11:34:59 | 4611 | Security System Extension |
|  | Überwachung gescheitert | 02.11.2025 11:34:59 | 4625 | Logon                     |
|  | Überwachung erfolgreich | 02.11.2025 11:34:59 | 5059 | Other System Events       |
|  | Überwachung erfolgreich | 02.11.2025 11:34:59 | 5059 | Other System Events       |
|  | Überwachung erfolgreich | 02.11.2025 11:34:59 | 4611 | Security System Extension |

# SIEM

## Security Information and Event Management: **Collection of logs + parsing**



# SIEM solutions

Splunk Enterprise Security

The logo for Splunk Enterprise Security, featuring the word "splunk" in black and ">enterprise" in green, all within a light gray rectangular background.

Microsoft Sentinel



IBM QRadar



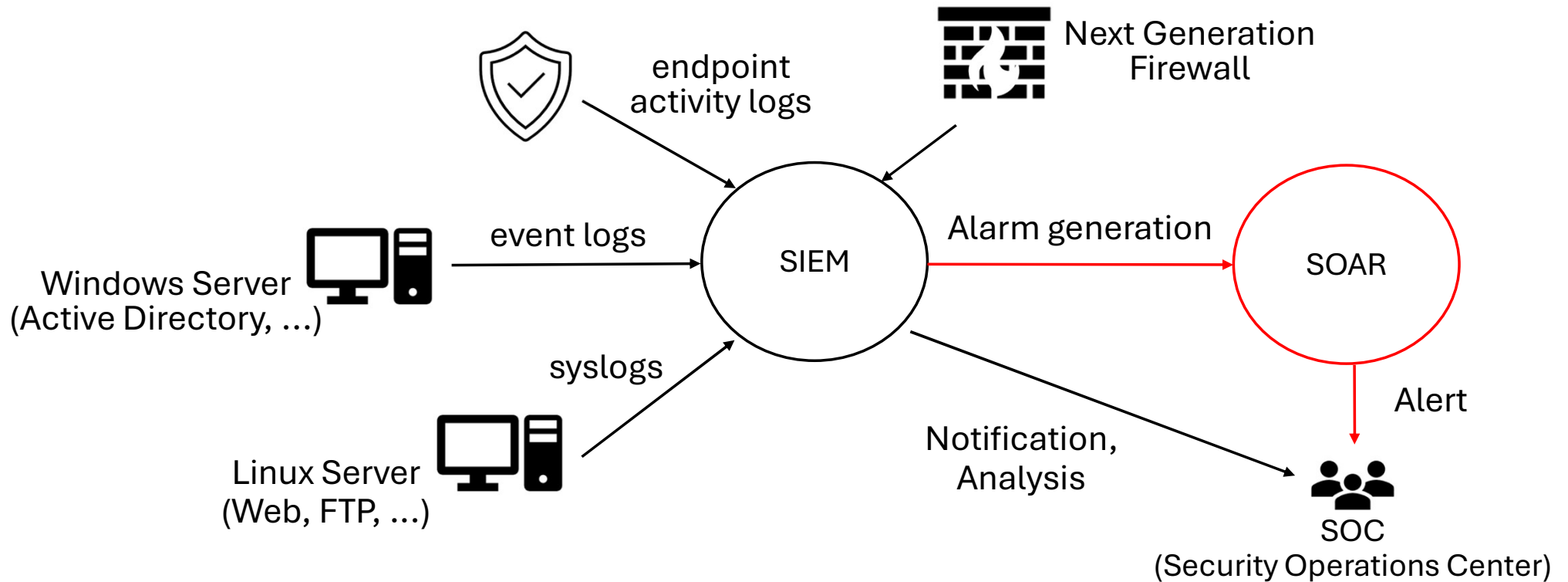
Elastic Security



Open source solutions (Elastic Stack, OSSIM, OSSEC, ...)

# SOAR

Security Orchestration, Automation, and Response:

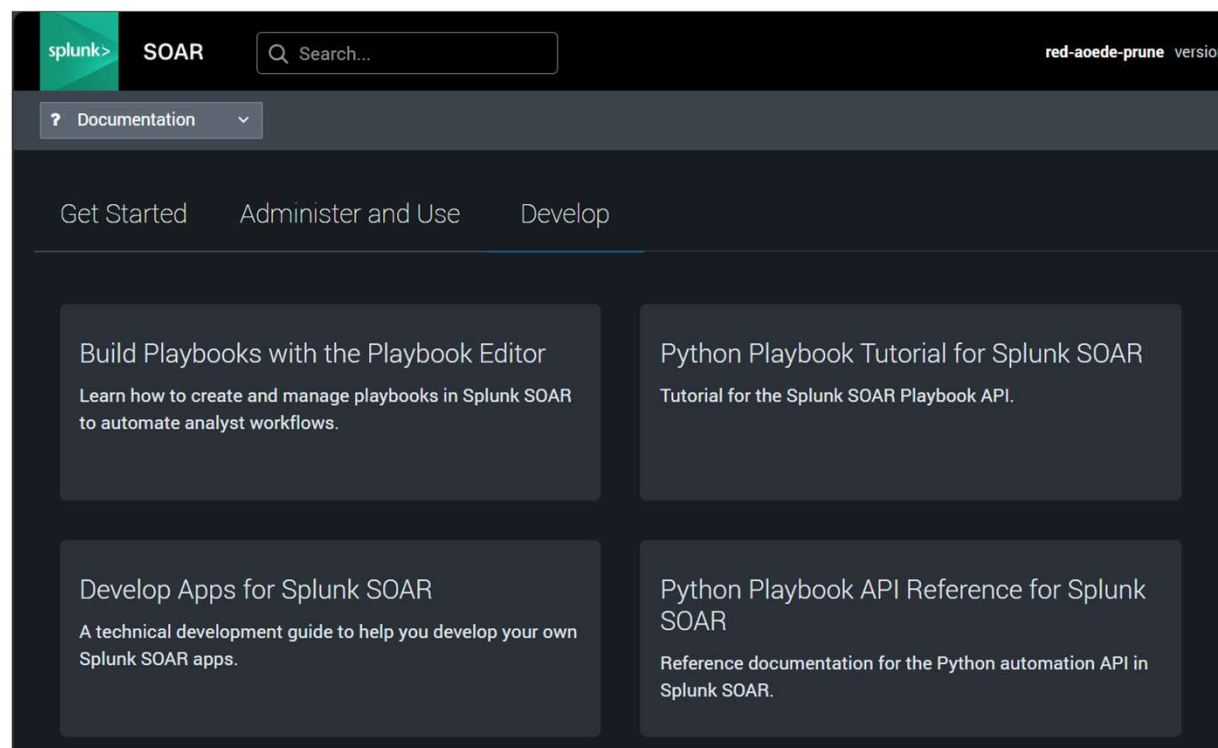




# Splunk SOAR

## Setup:

- On-premises installation
- Free Community Edition
- Phantom Community Playbooks (Apache-2.0 license)

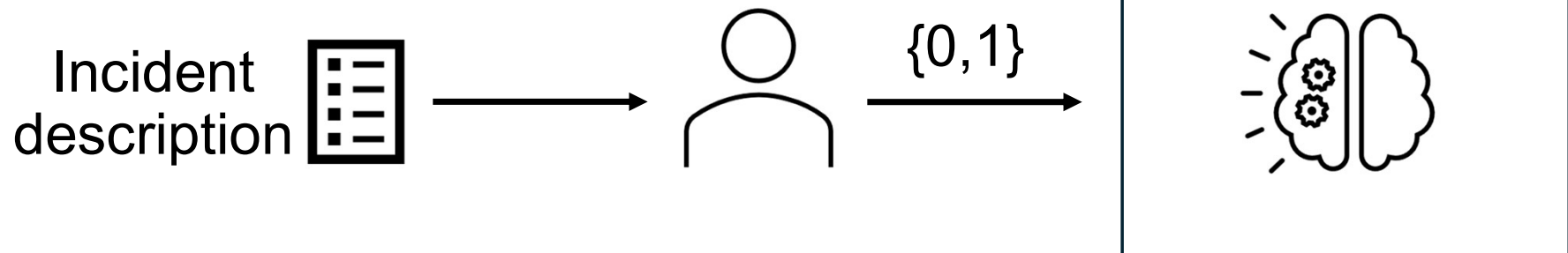


[https://www.splunk.com/en\\_us/download/soar-free-trial.html](https://www.splunk.com/en_us/download/soar-free-trial.html)

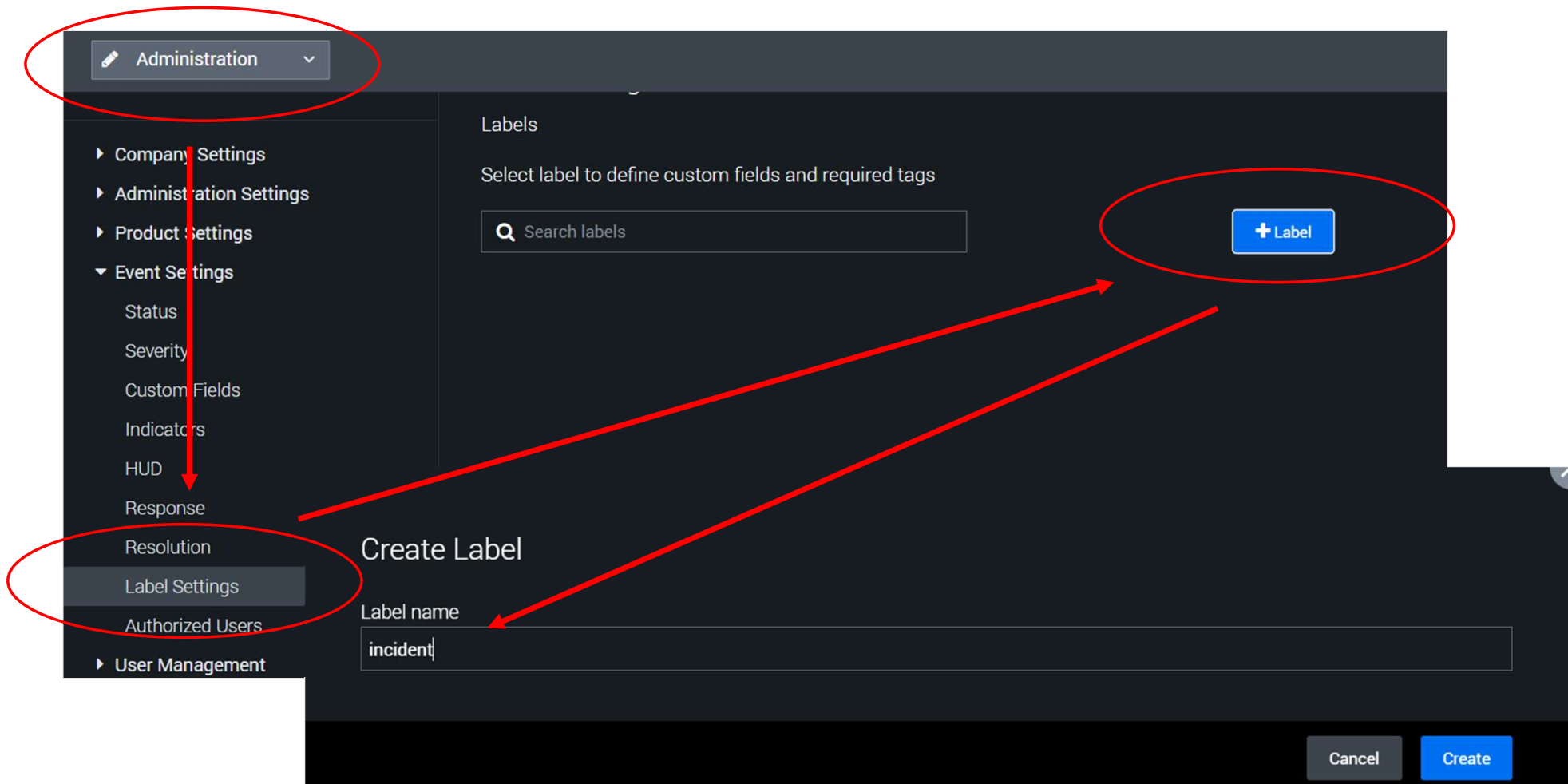
# How to reduce the FP

## Methodology

- Manual labelling → Ground Truth
- Training a Machine Learning (ML) model
- Continual Learning through human annotations



# Create new label



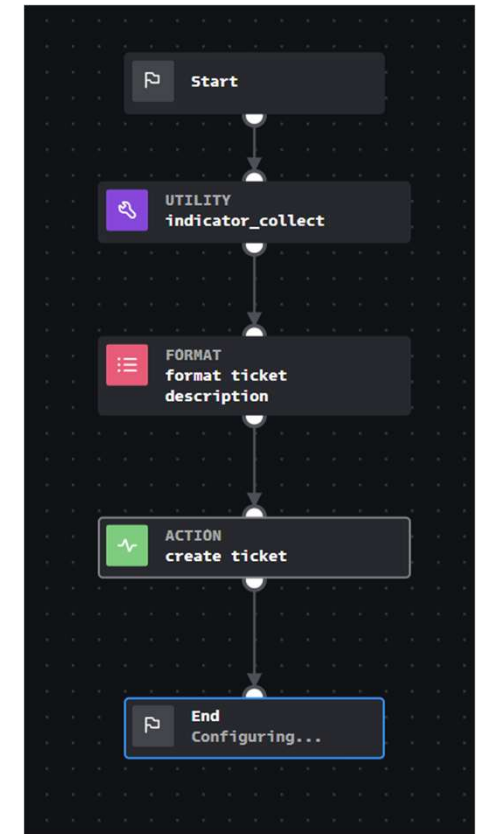
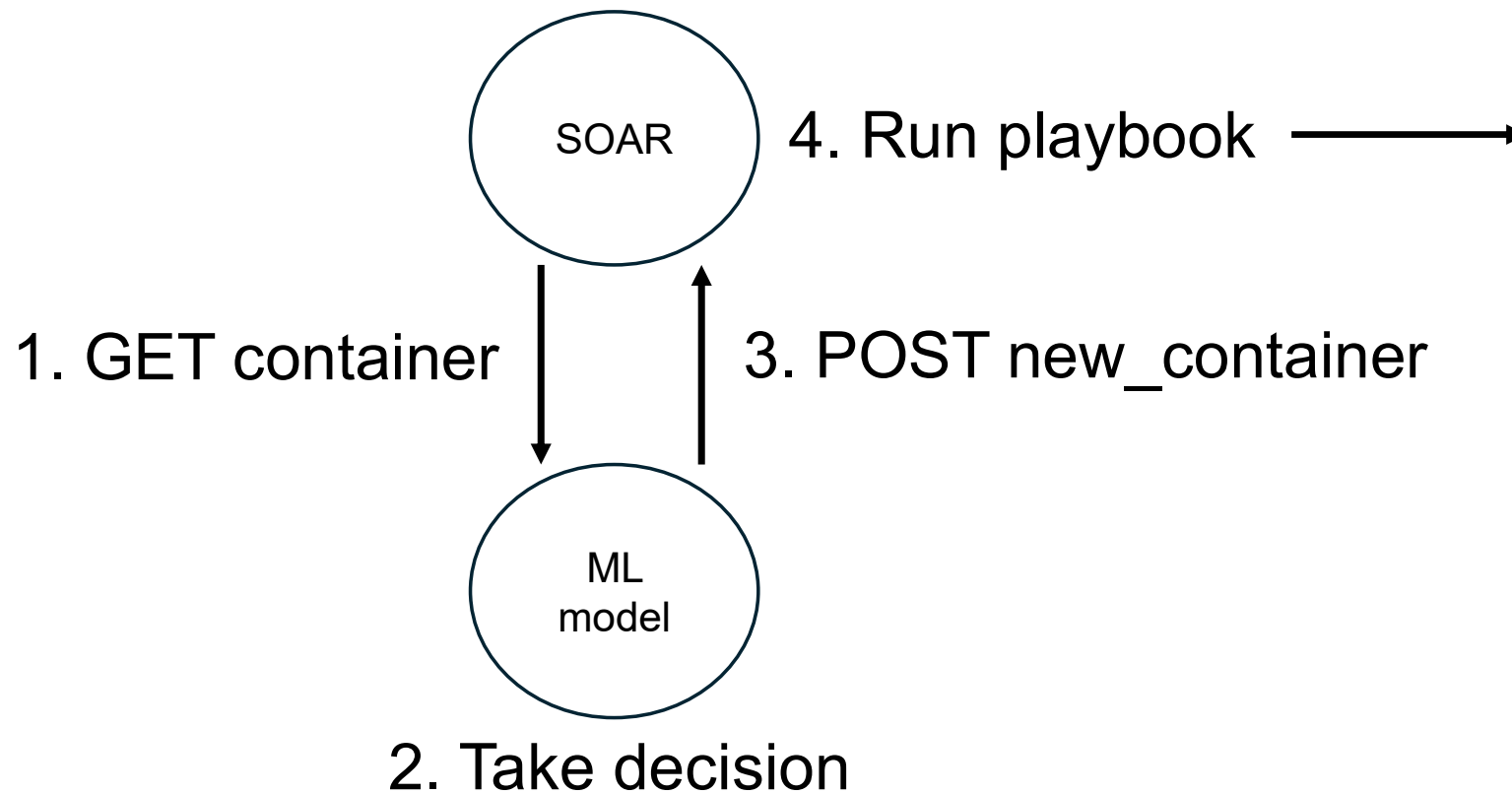
# Manual Labelling

The screenshot displays a security dashboard with a dark theme. At the top, a navigation bar includes a 'Cases' dropdown menu (circled in red) and tabs for 'Events', 'Cases', 'Tasks', and 'Indicators'. Below the navigation bar is a search bar labeled 'Search by event names or ID' and a 'Show' button. The main content area features three summary cards: 'Top Events' showing 7 events and 1 alert, 'Severity' with a donut chart and counts for High (3), Medium (4), and Low (1), and 'Status' with a donut chart and counts for New (8), Open (0), and Closed (0). At the bottom is a table of events. The first row of the table is circled in red, highlighting the 'NAME' column header and the event 'Manual Test Container'. A red arrow points from the 'Cases' dropdown menu to this circled row.

| ID | NAME                  | LABEL  | OWNER         | STATUS | SEVERITY |
|----|-----------------------|--------|---------------|--------|----------|
| 1  | Manual Test Container | events | Administrator | New    | LOW      |
| 2  | Test Container 1      | events | Automation    | New    | MEDIUM   |

# AI Integration

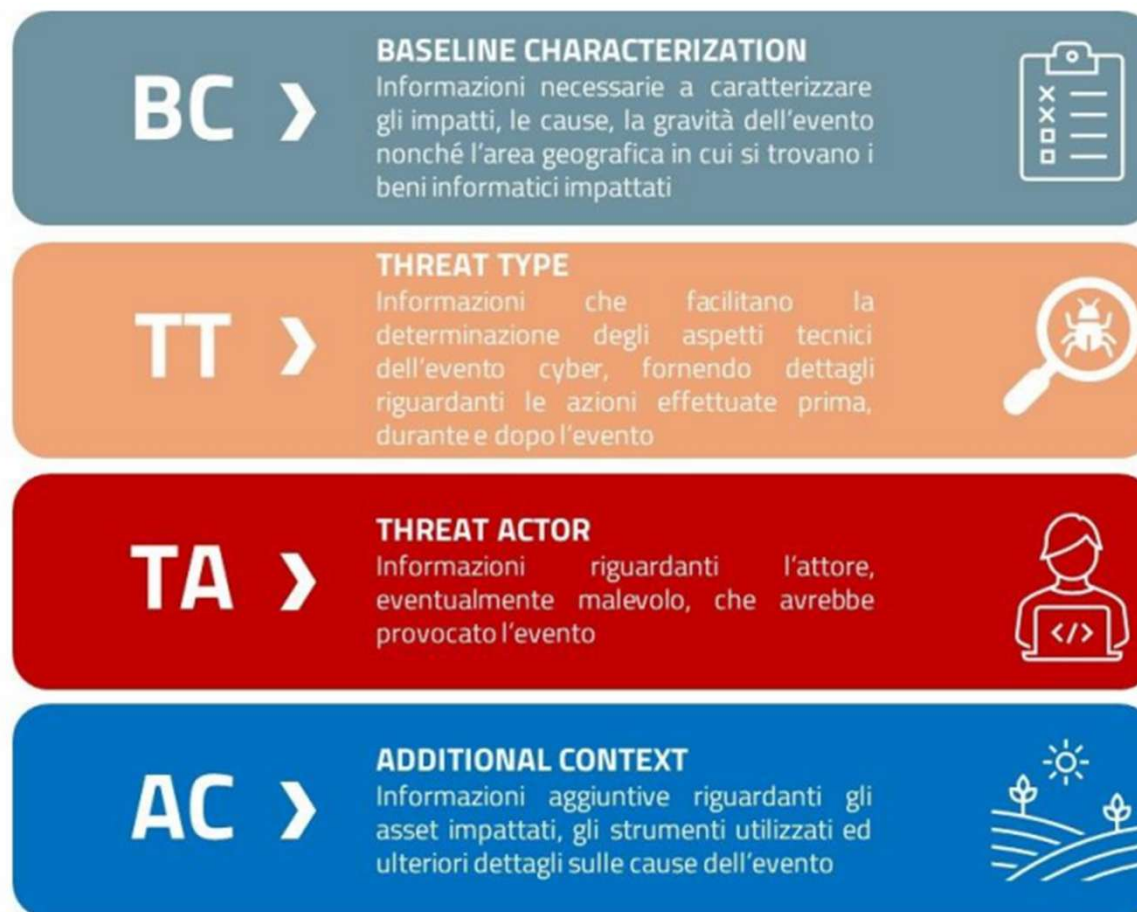
## Visual Playbook Editor



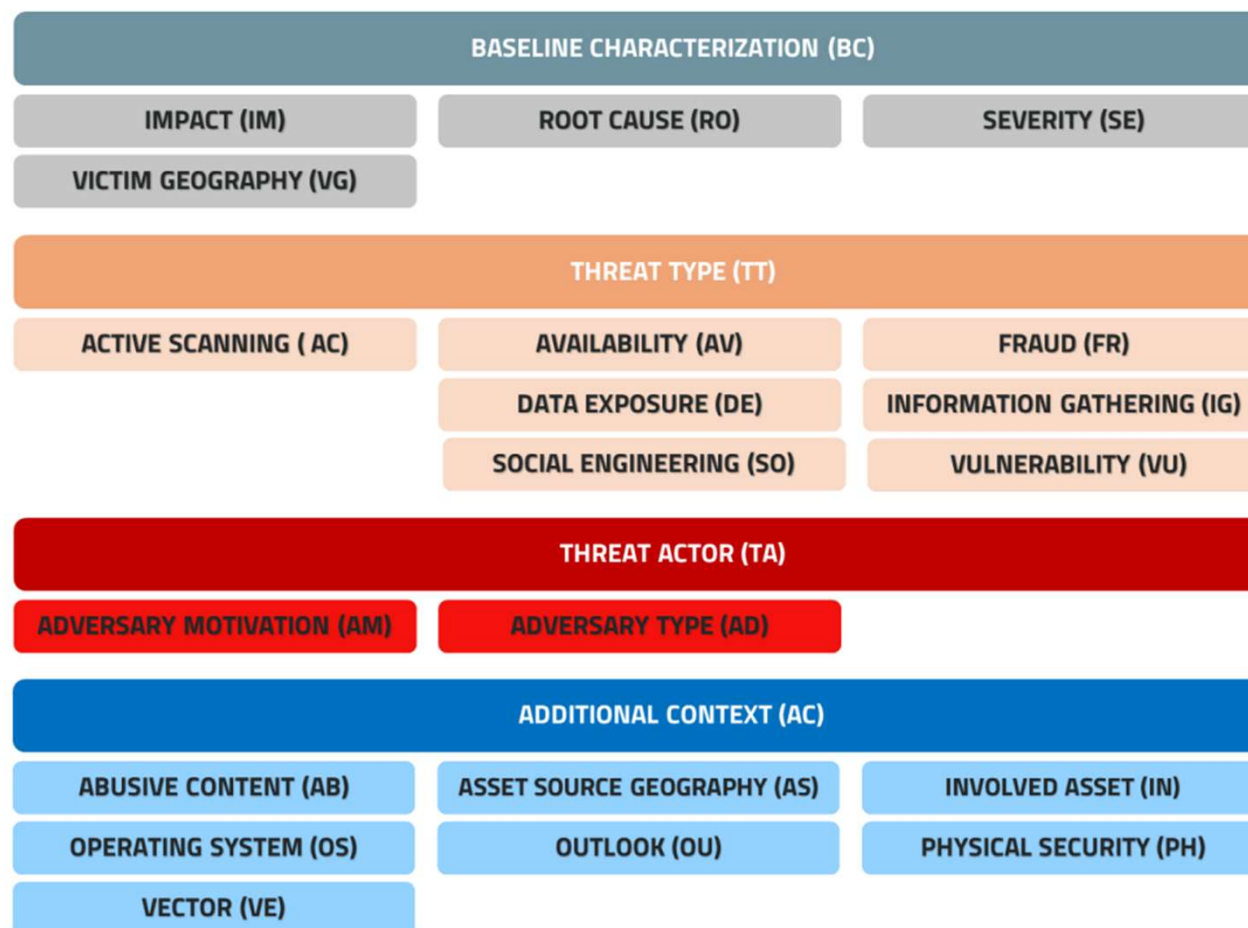
# Playbooks

- Phantom Community Playbooks
- Phantom library for Python
- Splunk playbooks: <https://research.splunk.com> for the Visual Playbook Editor (VPE)
- Phantom Community Playbooks on github: <https://github.com/phantomcyber/playbooks> (Apache-2.0 license)

# Incident reporting according to TC-ACN



# Macrocategories and Predicates



[https://www.acn.gov.it/portale/documents/d/guest/acn\\_tassonomia\\_cyber\\_clear](https://www.acn.gov.it/portale/documents/d/guest/acn_tassonomia_cyber_clear)



# Values

Values for the macrocategory **Baseline Characterization** and the predicate **Impact**:

| BASELINE CHARACTERIZATION (BC)   |                                      |                                 |
|----------------------------------|--------------------------------------|---------------------------------|
| IMPACT (IM)                      |                                      |                                 |
| Account compromise<br>(BC:IM-AC) | Application compromise<br>(BC:IM-AP) | Availability<br>(BC:IM-AV)      |
| Data exfiltration<br>(BC:IM-DX)  | Data exposure<br>(BC:IM-DE)          | Data manipulation<br>(BC:IM-DM) |
| No impact<br>(BC:IM-NO)          | System compromise<br>(BC:IM-SY)      | Other<br>(BC:IM-OT)             |

[https://www.acn.gov.it/portale/documents/d/guest/acn\\_tassonomia\\_cyber\\_clear](https://www.acn.gov.it/portale/documents/d/guest/acn_tassonomia_cyber_clear)

# Example of TC-ACN vector

**<BC:IM-DE BC:RO-MA BC:SE-LO BC:VG-EU  
TT:DE-PD TT:MA-BA TT:VU-SE TA:AM-ES TA:AD-CR  
AC:AB-OT AC:AS-IT AC:IN-OT AC:PH-UN>**

- Impact: Data Exposure
- Root Cause: Malicious Actions
- Severity: Low
- Victim Geography: Europe
- Data Exposure: Personal data
- Malicious Code: Backdoor

...

# MITRE ATT&CK Framework

250 adversary tactics and techniques  
(<https://attack.mitre.org>)

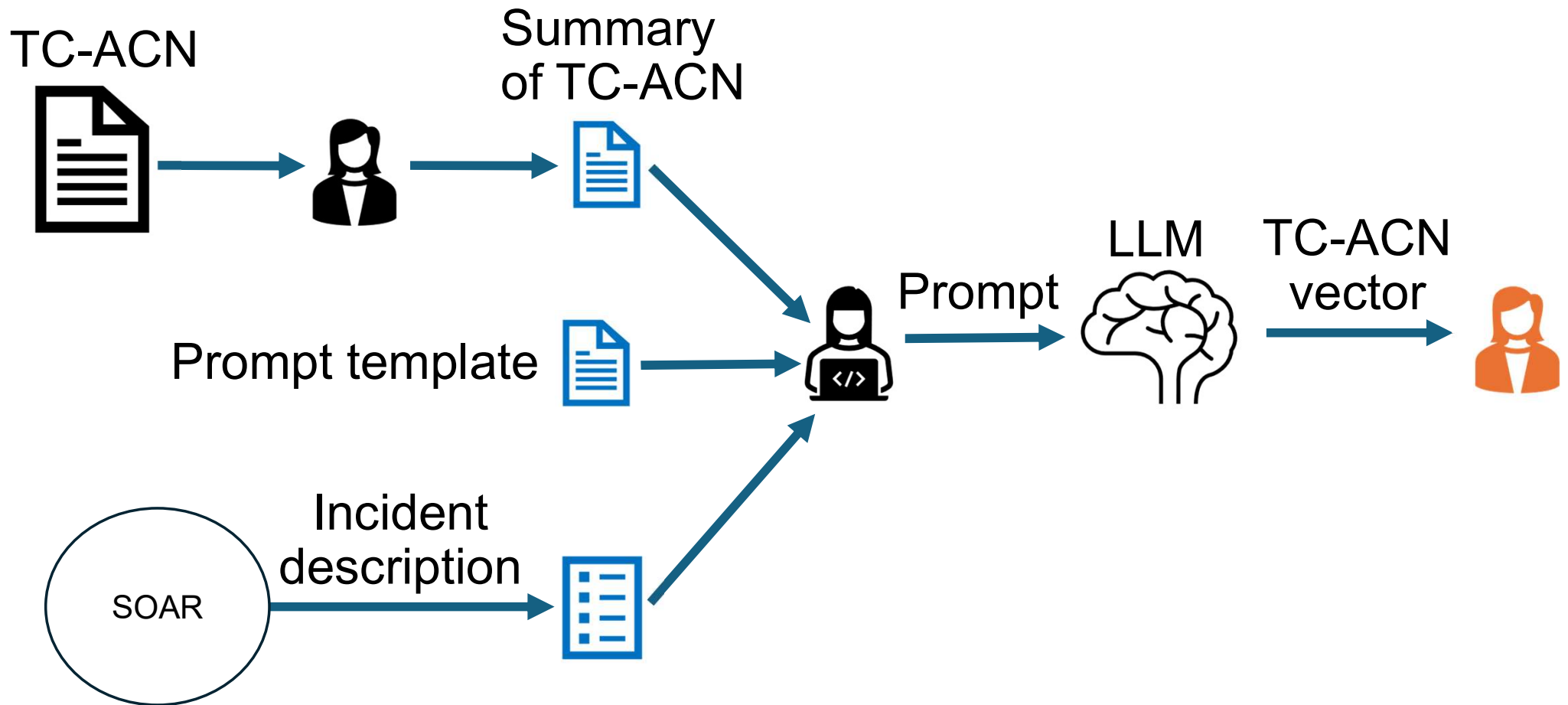
ATT&CK Matrix for Enterprise

layout: side ▾ show sub-techniques hide sub-techniques

| Reconnaissance                                                                                                                                                                                                                                                | Resource Development                                                                                                                                                                                                | Initial Access                                                                                                                                                                   | Execution                                                                                                                                                                                                            | Persistence                                                                                                                                                                                                | Privilege Escalation                                                                                                                                                                                                            | Defense Evasion                                                                                                                                                                                                                                                   | Credential Access                                                                                                                                                                                       | Discovery                                                                                                                                                                                                                                            | Lateral Movement                                                                                                                                                                         | Collection                                                                                                                                                                                | Command and Control                                                                                                                                                              | Exfiltration                                                                                                                                                                                                      | Impact                                                                                                                                                                                                   |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 10 techniques                                                                                                                                                                                                                                                 | 8 techniques                                                                                                                                                                                                        | 10 techniques                                                                                                                                                                    | 14 techniques                                                                                                                                                                                                        | 20 techniques                                                                                                                                                                                              | 14 techniques                                                                                                                                                                                                                   | 44 techniques                                                                                                                                                                                                                                                     | 17 techniques                                                                                                                                                                                           | 32 techniques                                                                                                                                                                                                                                        | 9 techniques                                                                                                                                                                             | 17 techniques                                                                                                                                                                             | 18 techniques                                                                                                                                                                    | 9 techniques                                                                                                                                                                                                      | 14 techniques                                                                                                                                                                                            |
| Active Scanning (3)<br>Gather Victim Host Information (4)<br>Gather Victim Identity Information (3)<br>Gather Victim Network Information (6)<br>Gather Victim Org Information (4)<br>Phishing for Information (4)<br>Search Closed Sources (2)<br>Search Open | Acquire Access<br>Acquire Infrastructure (8)<br>Compromise Accounts (3)<br>Compromise Infrastructure (8)<br>Develop Capabilities (4)<br>Establish Accounts (3)<br>Obtain Capabilities (7)<br>Stage Capabilities (4) | Content Injection<br>Drive-by Compromise<br>Exploit Public-Facing Application<br>External Remote Services<br>Hardware Additions<br>Phishing (4)<br>Replication Through Removable | Cloud Administration Command<br>Command and Scripting Interpreter (11)<br>Container Administration Command<br>Deploy Container<br>Exploitation for Client Execution<br>Inter-Process Communication (3)<br>Native API | Account Manipulation (7)<br>BITS Jobs<br>Boot or Logon Autostart Execution (14)<br>Boot or Logon Initialization Scripts (3)<br>Browser Extensions<br>Compromise Host Software Binary<br>Create Account (3) | Abuse Elevation Control Mechanism (6)<br>Access Token Manipulation (5)<br>Account Manipulation (7)<br>Boot or Logon Autostart Execution (14)<br>Boot or Logon Initialization Scripts (5)<br>Create or Modify System Process (3) | Abuse Elevation Control Mechanism (6)<br>Access Token Manipulation (5)<br>BITS Jobs<br>Build Image on Host<br>Debugger Evasion<br>Deobfuscate/Decode Files or Information<br>Deploy Container<br>Direct Volume Access<br>Domain or Tenant Policy Modification (2) | Adversary-in-the-Middle (4)<br>Brute Force (4)<br>Credentials from Password Stores (6)<br>Exploitation for Credential Access<br>Forced Authentication<br>Forge Web Credentials (2)<br>Input Capture (4) | Account Discovery (4)<br>Application Window Discovery<br>Browser Information Discovery<br>Cloud Infrastructure Discovery<br>Cloud Service Dashboard<br>Cloud Service Discovery<br>Cloud Storage Object Discovery<br>Container and Resource Discovery | Exploitation of Remote Services<br>Internal Spearphishing<br>Lateral Tool Transfer<br>Remote Service Session Hijacking (2)<br>Remote Services (8)<br>Replication Through Removable Media | Adversary-in-the-Middle (4)<br>Archive Collected Data (3)<br>Audio Capture<br>Automated Collection<br>Browser Session Hijacking<br>Clipboard Data<br>Data from Cloud Storage<br>Data from | Application Layer Protocol (3)<br>Communication Through Removable Media<br>Content Injection<br>Data Encoding (2)<br>Data Obfuscation (3)<br>Dynamic Resolution (3)<br>Encrypted | Automated Exfiltration (1)<br>Data Transfer Size Limits<br>Exfiltration Over Alternative Protocol (3)<br>Exfiltration Over C2 Channel<br>Exfiltration Over Other Network Medium (1)<br>Exfiltration Over Physical | Account Access Removal<br>Data Destruction (1)<br>Data Encrypted for Impact<br>Data Manipulation (3)<br>Defacement (2)<br>Disk Wipe (2)<br>Endpoint Denial of Service (4)<br>Financial Theft<br>Firmware |

<https://attack.mitre.org>

# Methodology



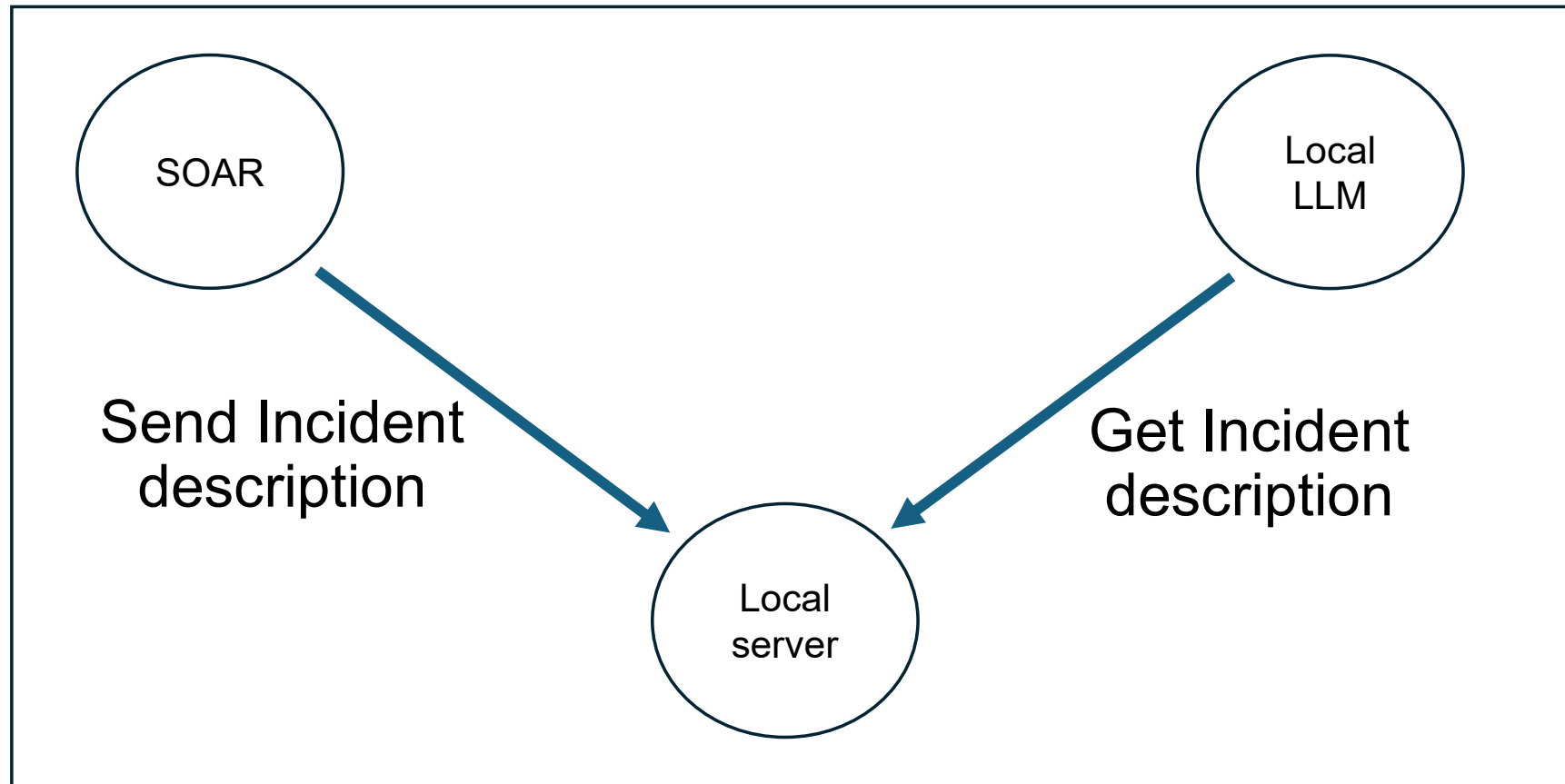
# How to generate the prompt

## **Prompt:**

You are part of a SOC team. Your role is the security specialist. Consider the incident description <incident description>. Generate the TC-ACN vector in the correct format according to the definition of the TC-ACN taxonomy <TC-ACN summary>.

# Automation

## On-premises solution



# Preliminary Results

**Setup:** gpt-oss:20b model, thinking feature, 4k context length

**Synthetic data** for the incident description

**Incident description:** „Potential DoS Attack detected“ →

**<BC:IM-AV BC:RO-MA BC:SE-ME BC:VG-GL TT:DS-OT  
AC:AB-OT AC:AS-OT AC:IN-OT AC:PH-OT TA:CY-OT>**

# Preliminary Results

**Setup:** gpt-oss:20b model, thinking feature, 4k context length

**Synthetic data** for the incident description

**Incident description:** „Web application exploited through SQL injection“ →

**<BC:IM-AP BC:RO-MA BC:SE-ME BC:VG-GL TT:VU-OT  
TA:CY-OT AC:AV-UN AC:TS-WS AC:ET-UN AC:DM-UN  
AC:RA-UN>**



# Conclusions

- AI Augmentation to address the problem of FP (it needs a ground truth / human feedback)
- Phantom Community Playbooks to test automated responses (basic concepts can be applied to different software solutions)
- Brief introduction to TC-ACN
- Leveraging (local) LLMs
- Bias problem → Improving the quality of the TC-ACN summary
- Introducing AI models to replace humans
- Fine-Tuning LLMs with examples



Finanziato  
dall'Unione europea  
NextGenerationEU



Ministero  
dell'Università  
e della Ricerca



Italiadomani  
PIANO NAZIONALE  
DI RIPRESA E RESILIENZA



This work is co-funded by the Italian MIUR (with NRRP funds) and by the local company SIAG / Informatica Alto Adige S.p.A. with the scholarship provided under the project number CUP I52B24000510005.

We would like to acknowledge the assistance and the support of Ing. Francesco Terracciano and Mr. Abdalrahman Hwoij from the company SIAG / Informatica Alto Adige S.p.A.

# References and Links

- Agenzia per la cybersicurezza nazionale (2025), La direttiva NIS, <https://www.acn.gov.it/portale/nis>
- Agenzia per la cybersicurezza nazionale (2025), La Tassonomia Cyber dell'ACN, [https://www.acn.gov.it/portale/documents/d/guest/acn\\_tassonomia\\_cyber\\_clear](https://www.acn.gov.it/portale/documents/d/guest/acn_tassonomia_cyber_clear)
- Ollama (2025), Ollama models, <https://ollama.com/search>
- Splunk playbooks (2025): <https://research.splunk.com>
- Phantom Community Playbooks on github (2025): <https://github.com/phantomcyber/playbooks>
- MITRE ATT&CK Framework (2025): <https://attack.mitre.org>

Note: All listed links were last accessed on 02/11/2025.